UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/019,882	04/15/2002	Yan Yonghong	42390.P8351	8063

7590 09/26/2008

Blakely Sokoloff Taylor& Zafman
12400 Wilshire Boulevard
Los Angeles, CA 90025

EXAMINER	
WOZNIAK, JAMES S	

ART UNIT	PAPER NUMBER
2626	

MAIL DATE	DELIVERY MODE
09/26/2008	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No.	Applicant(s)	
	10/019,882	YONGHONG, YAN	
	Examiner	Art Unit	
	JAMES S. WOZNIAK	2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *17 June 2008*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-30* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☒ Claim(s) *4,12,19 and 27* is/are allowed.

6)☒ Claim(s) *1-3,5,7-11,13,15-18,20,22-26,28 and 30* is/are rejected.

7)☒ Claim(s) *6,14,21 and 29* is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *15 April 2002* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

Response to Amendment

1. In response to the office action from 3/18/2008, the applicant has submitted an amendment, filed 6/17/2008, amending independent claims 16, 19, 23, and 27, while arguing to traverse the art rejection based on the limitations regarding calculating estimated weights, marking utterance sections, and using weighted sections for speaker independent-dependent model conversion *(Amendment, Pages 10-11)*. Applicant's arguments have been fully considered, however the previous rejection is maintained due to the reasons listed below in the response to arguments.

2. In response to the amendment of claims 16, 19, 23, and 27 *(Amendment, Page 9)*, which includes that the instructions are executable and executed by a processor to realize the practical application functionality of the presently claimed invention, the examiner has withdrawn the previous 35 U.S.C. 101 rejection.

Response to Arguments

3. Applicant's arguments have been fully considered but they are not persuasive for the following reasons:

With respect to the independent claims *(1, 8, 16, and 23)*, the applicant argues that Barnard et al *(U.S. Patent: 7,216,079)* first fails to teach calculated estimated weights for an identified error in recognition of utterances based on a reference string because the segment alignment data in Barnard only represents boundaries and there is no teaching of a reference string *(Amendment, Page 10)*.

In response, the examiner points out that while Barnard does utilize alignment data in detecting speech recognition errors for model training, the alignment data is not what was relied upon in the previous Office Action for teaching the "estimated weights". Instead, it is the determined shifting amount that is applied to mean vectors of acoustic models for training that anticipates this claimed feature *(Col. 3, Line 64- Col. 4, Line 11; and Col. 6, Lines 21-39)*. This shifting amount represents a numerical modification to the mean values of a speech recognition model, and thus, effectively weights an acoustic model numerically in one direction or another *(i.e., closer or further)* based on a correct/incorrect speech recognition decision *(Col. 3, Line 64-Col. 4, Line 11)*. Also, Barnard's comparison involves a cross-comparison involving a correct phoneme string sequence that corresponds to a word. It is through comparison with this reference string that recognition errors are determined *(Col. 6, Lines 21-39)*. Thus, Barnard does teach a reference string. Therefore, for at least the above reasons, the applicant's first argument has been fully considered, but is not convincing.

With respect to the independent claims, the applicant secondly argues that Barnard merely discloses alignment between different segmentations and not marking sections as being misrecognized and further alleges that since there are no weights, there can be no associating of weights and sections *(Amendment, Page 10)*.

In response, the examiner notes that Barnard does teach marking sections as being misrecognized. In his comparison, Barnard matches up a second alignment with a correct, reference sequence string. As Barnard proceeds through this comparison, wrong phonemes and correct phonemes within the second sequence are marked with respect to the reference string *("correct and wrong", "incorrectly recognized", Col. 6, Lines 21-39).* Thus, since Barnard sequentially compares a second utterance with a reference to determine those phoneme sections that are wrong or "incorrectly recognized", Barnard does anticipate the section marking recited in the presently claimed invention. Furthermore, the shifting values or weights in Barnard are assigned to the sections based on this marking *(Col. 3, Line 64- Col. 4, Line 5).* Thus, the applicant's second argument has been fully considered, but is not convincing.

Finally, with respect to the independent claims, the applicant argues that Barnard only teaches moving incorrect phonemes away from a mean value and does not teach using weighted utterance sections to convert a speaker independent model to a speaker dependent model *(Amendment, Pages 10-11).*

In response, the examiner notes that in Barnard's process is directed to training a speech recognition model *(Col. 3, Line 64- Col. 4, Line 11).* The data used to train/modify the initial system model is based on utterances received from a particular application user *(Col. 6, Lines 40-46).* An initial, untrained correct acoustic model in Barnard would not involve any received user speech data, but is trained over time according to speech received from a user. In this way, the training process of Barnard begins with a non-user based acoustic model and ends in a user-trained or speaker dependent model *(i.e., it converts SI to SD).* Thus, for at least this reason, the applicant's last argument has been fully considered, but is not convincing.

The applicant's arguments with respect to Claim 5 *(Amendment, Page 13)* are identical to those presented in the response filed on 12/26/2007. In the subsequent Office Action from 3/18/2008 *(Pages 3-4)*, the examiner explained why the two equations are equivalent and that motivation was provided for including the teachings of Junqua *(U.S. Patent: 6,253,181)*. Since these arguments are similar to those presented previously and since the applicant has not specifically addressed the corresponding examiner response, please see Pages 3-4 of the Office Action from 3/18/2008 in regards to these arguments.

The art rejections of the respective dependent claims are traversed for reasons similar to the independent claims *(i.e., 1, 8, 16, and 23) (Amendment, Page 14)*. In regards to such arguments see the above response directed towards the independent claims.

Claim Rejections - 35 USC § 102

4. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

5. **Claims 1-3, 7-11, 15-18, 22-26, and 30** are rejected under 35 U.S.C. 102(e) as being anticipated by Barnard et al *(U.S. Patent: 7,216,079)*.

With respect to **Claim 1**, Barnard discloses:

Calculating estimated weights for identified errors in recognition of utterances based on a reference string *(using reference strings to identify incorrectly recognized utterance sections and determining associated training weights, Col. 3, Line 64- Col. 4, Line 11; Col. 5, Lines 16-26; Col. 6, Lines 21-39; and Col. 9, Lines 47-67; and Fig. 3)*;

Marking sections of the utterances as being misrecognized and associating the estimated weights with the sections of the utterances *(utterance segment locations that are incorrectly recognized are selected and associated with a training weight shift value, Col. 6, Lines 21-39; Col. 8, Lines 51-60; Col. 9, Lines 47-67; and Fig. 3)*;

Using the weighted sections of the utterances to convert a speaker independent model to a speaker dependent model *(weighted utterance segments are used to gradually train an initial model for a particular speaker, Col. 3, Line 64- Col. 4, Line 11; and Col. 6, Lines 40-61)*.

With respect to **Claim 2**, Barnard further discloses:

The method steps (a)-(c) are repeated at least once *(repeated processing is performed, Col. 3, Line 64- Col. 4, Line 11; and Col. 6, Lines 40-46)*.

With respect to **Claim 3**, Barnard further discloses:

The utterances are converted into a recognized phone string a first time through applying the speaker independent model and thereafter through applying the most recently obtained speaker dependent model *(recognizer creates phoneme strings using an initial model that is gradually/repeatedly trained, Col. 3, Line 64- Col. 4, Line 11; and Col. 6, Lines 40-46; and Fig. 3)*.

With respect to **Claim 7**, Barnard further discloses:

Different misrecognized words have different weights *(variable weighting value, Col. 9, Lines 47-67)*.

With respect to **Claim 8**, Barnard discloses:

Recognizing utterances through converting the utterances into a recognized string *(speech recognition generates a phoneme string, Col. 7, Lines 13-46)*;

Comparing the recognized string with a reference string to determine errors *(location of errors is determined by comparing correct reference string and recognized string, Col. 8, Lines 51-60)*;

Calculating estimated weights for sections of the utterances *(using reference strings to identify incorrectly recognized utterance sections and determining associated training weights, Col. 3, Line 64- Col. 4, Line 11; Col. 5, Lines 16-26; Col. 6, Lines 21-39; and Col. 9, Lines 47-67; and Fig. 3)*;

Marking the errors in the utterances and providing corresponding estimated weights to form adaptation enrollment data *(utterance segment locations that are incorrectly recognized are selected and associated with a training weight shift value, Col. 6, Lines 21-39; Col. 8, Lines 51-60; Col. 9, Lines 47-67; and Fig. 3)*; and

Using the adaptation enrollment data to convert a speaker independent model to a speaker dependent model *(weighted utterance segments are used to gradually train an initial model for a particular speaker, Col. 3, Line 64- Col. 4, Line 11; and Col. 6, Lines 40-61)*.

With respect to **Claim 9**, Barnard further discloses:

The utterances are converted into the recognized string through applying the speaker independent model *(initial recognition model that is to be gradually adapted, Col. 3, Line 64- Col. 4, Line 15)*.

With respect to **Claim 10**, Barnard further discloses:

Parts (b)-(e) are repeated until differences between the reference and recognized strings are less than a threshold (corrective action is only taken until a difference greater than a closeness threshold *(i.e., below an effective threshold measure of similarity) is reached, Col. 9, Lines 11-26; and Col. 10, Lines 14-22)*.

Claim 11 contains subject matter similar to Claim 3, and thus, is rejected for the same reasons.

Claim 15 contains subject matter similar to Claim 7, and thus, is rejected for the same reasons.

With respect to **Claim 16**, Barnard discloses the method for marking and weighting misrecognized utterance sections for speaker training as applied to claim 1, implemented as a computer readable medium storing a program executable by a computer *(Col. 11, Line 49- Col. 12, Line 18)*.

Claims 17-18 contain subject matter respectively similar to Claims 2-3, and thus, are rejected for the same reasons.

Claim 22 contains subject matter similar to Claim 7, and thus, is rejected for the same reasons.

With respect to **Claim 23**, Barnard discloses the method for marking and weighting misrecognized utterance sections for speaker training as applied to claim 8, implemented as a

computer readable medium storing a program executable by a computer *(Col. 11, Line 49- Col. 12, Line 18)*.

Claims 24-26 contain subject matter respectively similar to Claims 9-11, and thus, are rejected for the same reasons.

Claim 30 contains subject matter similar to Claim 7, and thus, is rejected for the same reasons.

Claim Rejections - 35 USC § 103

6. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

7. **Claims 5, 13, 20, and 28** are rejected under 35 U.S.C. 103(a) as being unpatentable over Barnard et al in view of Junqua *(U.S. Patent: 6,253,181)*.

With respect to **Claims 5, 13, 20, and 28**, Barnard discloses the method for marking and weighting misrecognized utterance sections for speaker training, as applied to Claims 1, 8, 16, and 23. Nguyen does not specifically disclose that calculation of a weighting score that computes an average likelihood difference per frame, however Junqua discloses a calculation of a likelihood difference used in determining a speaker adaptation that utilizes an average of likelihood difference scores associated with an incorrect recognition *(Col. 4, Lines 9-24; and Col. 5, Lines 15-67)*. Junqua further discloses an equation similar to that recited in claim 5 for

determining a log-likelihood difference in a speaker adaptation process that utilizes an average of likelihood scores *(Col. 5, Lines 15-67; and Col. 4, Lines 9-24)*.

Barnard and Junqua are analogous art because they are from a similar field of endeavor in speaker adaptation systems. Thus, it would have been obvious to a person of ordinary skill in the art, at the time of invention, to modify the teachings of Barnard with the likelihood difference calculation taught by Junqua in order to implement a high speed speaker adaptation system that is capable of providing a measure of recognition reliability *(Junqua, Col. 3, Lines 29-31; and Col.4, Lines 9-24)*.

Allowable Subject Matter

8. **Claims 4, 12, 19, and 27** are allowable over the prior art of record.

9. The following is an examiner's statement of reasons for allowance:

With respect to **Claims 4, 12, 19, and 27**, the prior art of record fails to explicitly teach or fairly suggest a method or computer readable medium storing a program executed by a computer for speaker adaptation that utilizes estimated weights based on misrecognized speech utterances as respectively recited in claims 4 and 12, wherein the estimated weights are calculated by computing an average likelihood difference per frame and then computing a weight value by averaging the average likelihood difference over error words (specification, page 6).

Although Barnard et al *(U.S. Patent: 7,216,079)* discloses that it is well known in the prior art to mark and weight misrecognized utterance sections for speaker training *(Col. 3, Line*

64- Col. 4, Line 11; Col. 5, Lines 16-26; Col. 6, Lines 21-39; and Col. 9, Lines 47-67; and Fig. 3) and Junqua *(U.S. Patent: 6,253,181)* teaches an equation for calculating an average likelihood difference, as applied to claim 5, Junqua does not teach averaging the average likelihood difference over all error words to determine a weight for speaker adaptation of a speech recognition model. Thus, claims 4 and 12 are allowable over the prior art of record.

Any comments considered necessary by applicant must be submitted no later than the payment of the issue fee and, to avoid processing delays, should preferably accompany the issue fee. Such submissions should be clearly labeled "Comments on Statement of Reasons for Allowance."

10. **Claims 6, 14, 21, and 29** are objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims.

11. The following is a statement of reasons for the indication of allowable subject matter:

With respect to **Claims 6, 14, 21, and 29**, the prior art of record fails to explicitly teach or fairly suggest a method for speaker adaptation that utilizes estimated weights based on misrecognized speech utterances, wherein the estimated weights are calculated by multiplying an average likelihood difference per frame calculated using the equation recited in claims 5, 13, 20, and 28 by the inverse of a number of misrecognized words for a particular speaker as per the equation recited in claims 6, 14, 21, and 29.

Although Barnard et al *(U.S. Patent: 7,216,079)* discloses that it is well known in the prior art to mark and weight misrecognized utterance sections for speaker training *(Col. 3, Line 64- Col. 4, Line 11; Col. 5, Lines 16-26; Col. 6, Lines 21-39; and Col. 9, Lines 47-67; and Fig. 3)* and Junqua *(U.S. Patent: 6,253,181)* teaches an equation for calculating an average likelihood difference, Junqua does not teach multiplying the calculated average likelihood by the inverse of a number of misrecognized words for a particular speaker as per the equation recited in claims 6, 14, 21, and 29.

Conclusion

12. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure: See PTO-892.

13. Any inquiry concerning this communication or earlier communications from the examiner should be directed to James S. Wozniak whose telephone number is (571) 272-7632. The examiner can normally be reached on M-Th, 7:30-5:00, F, 7:30-4, Off Alternate Fridays.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Patrick Edouard can be reached at (571) 272-7603. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

/James S. Wozniak/
Patent Examiner, Art Unit 2626